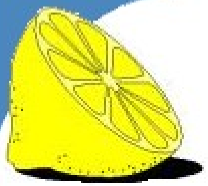


# Discovering Fuzzy Knowledge from Data

Tampere, Finland
Friday, November 2nd 2007

© Ján Boháčik
janbohacik@yahoo.com

# Outline

- ❑ Introduction
  - ✓ Knowledge and fuzzy knowledge
  - ✓ Fuzzy logics
  - ✓ Fuzzy rule base
- ❑ Mining Fuzzy Rules
  - ✓ Algorithm Based on a Fuzzy Decision Tree
  - ✓ Method Based on Variable Elimination
  - ✓ SW Tool Rule Miner
  - ✓ Algorithm Comparison
- ❑ Discussion

janbohacik@yahoo.com

# City of Zilina

- ❑ Zilina (Slovakia)
  - ✓ 156 000 inhabitants in the region
  - ✓ 86 000 inhabitants in the city
- ❑ Industry
  - ✓ Car industry
  - ✓ Heavy industry
  - ✓ Wood processing
- ❑ Companies
  - ✓ Kia, Mobis and suppliers
  - ✓ Siemens
  - ✓ Scheid & Bachman

# University of Zilina

## ❑ History:

- 1953 – established as a College of Railways in Prague
- 1959 – renamed the University of Transport
- 1980 – moved to Zilina
- 1996 – University of Zilina

## ❑ Main R&D and educational interests:

- Transportation and logistic
- Telecommunications
- Economy and management
- Military logistic and engineering
- Civil engineering
- Forensic engineering

janbohacik@yahoo.com

# Faculty of Management and Informatics

- Bachelor degree programs:
  - Informatics (Computer Science)
  - Computer Engineering
  - Management
- Master degree programs:
  - Information Systems
  - Economic Informatics
  - Computer Engineering
  - Information Management
- PhD programs:
  - Applied Informatics
  - Management
  - Control and Transportation Systems

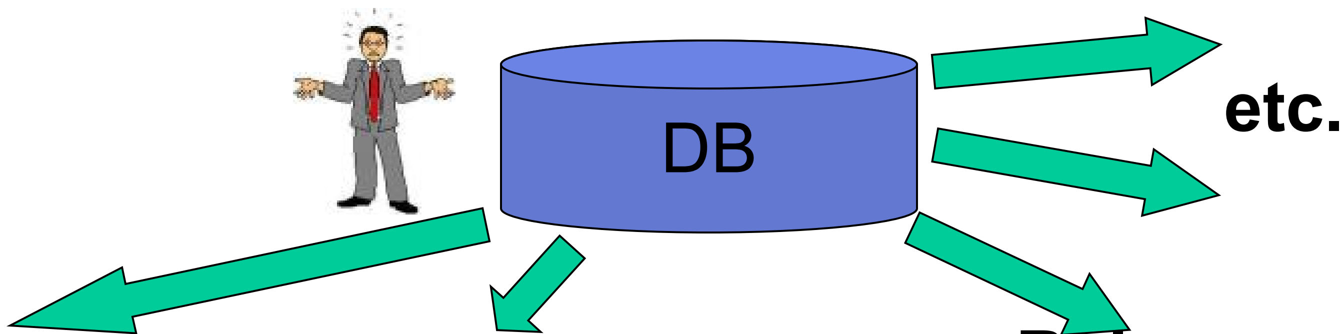
janbohacik@yahoo.com

# Department of Informatics

- Comprises about 20 academics and researchers

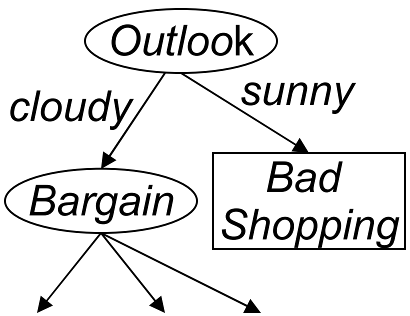
- Research and projects in:
  - Distributed and Parallel Systems
  - Reliability Analysis
  - Decision Making Support Systems Based on Fuzzy Logics, Data Mining
  - Intelligent Transportation and Environment Systems
  - Location Based Services
  - E-Learning and E-Payment Technologies

# What Is Knowledge?




**Decision Tree** | **Neural Network** | **Rules**

## Decision Tree




- Only classification
- Important attributes
- Into rules
- Categorical data

## Neural Network

- Classification, prediction
- Non-understandable results
- Overfitting
- Numerical data

## Rules

**IF** *Outlook* is *sunny* **THEN** *Shopping* is *bad*

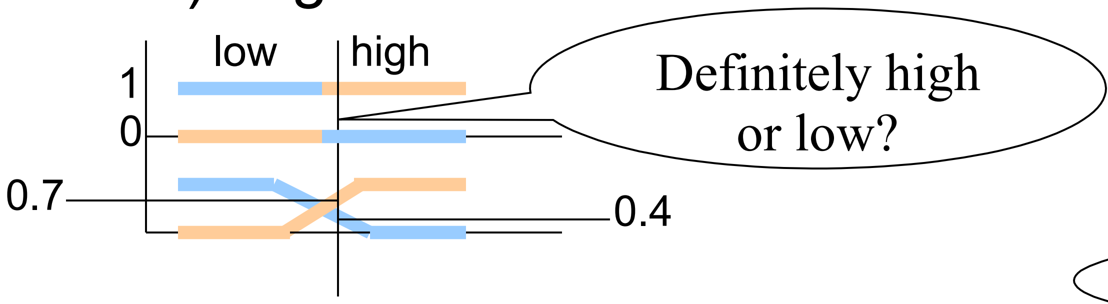
**IF** *Outlook* is *cloudy* **AND** *Bargain* is *none* **THEN** *Shopping* is *good*

- Classification, association
- Close to human thinking
- Makable and Usable by both humans and computers
- Lots of interpretations
- Basically categorical data

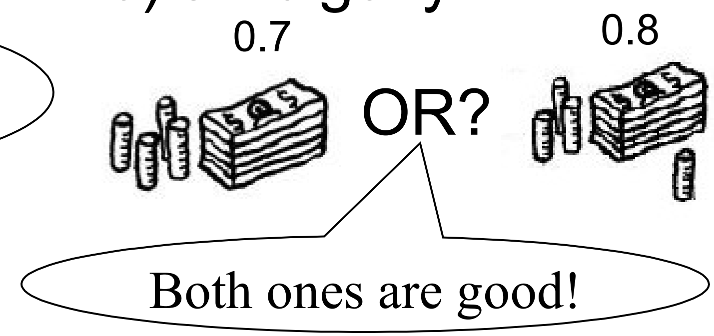
# Why Fuzzy Knowledge?

•Cognitive uncertainties [Klir, 1987]

a) vagueness




b) ambiguity




- small changes in categorical attribute values can cause rapid and inadequate changes in classes [Quinlan, 1987]

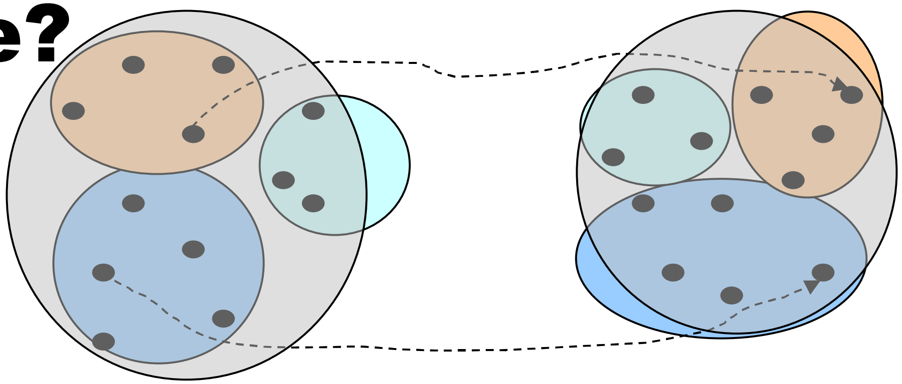
**[Klir, 1987]** Klir, G. J.: Where do we stand on measures of uncertainty, ambiguity, fuzziness and the like? *Fuzzy Sets and Systems* 24, 1987, 141-160.

**[Quinlan, 1987]** Quinlan, J. R.: Decision trees at probabilistic classifiers. *In Proceedings of the 4th International Workshop on Machine Learning*, CA, 1987, 31-37.

janbohacik@yahoo.com

# Why Fuzzy Knowledge?

- Reduction of complexity
  - fewer attributes and possible values
  - keeping knowledge accuracy




| *Instance* | *Tempera-ture* | *Outlook* | | | *Novelty* | *Shopping* |
|---|---|---|---|---|---|---|
| | | *rainy* | *overcast* | *sunny* | | |
| ($e_1$) | 15,4543 | 0,1 | 0,8 | 0,1 | *yes* | 20 000 |
| ($e_2$) | 16,0138 | 0,7 | 0,3 | 0,0 | *no* | 15 000 |
| ($e_3$) | 11,9943 | 0,9 | 0,1 | 0,0 | *yes* | 14 000 |
| ($e_4$) | 14,2764 | 0,2 | 0,2 | 0,6 | *no* | 14 800 |

| *Instance* | *Temperature* | | | *Outlook* | | | *Novelty* | | *Shopping* | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *cool* | *mild* | *hot* | *rainy* | *overcast* | *sunny* | *no* | *yes* | *good* | *bad* |
| ($e_1$) | 0,0 | 0,9 | 0,1 | 0,1 | 0,8 | 0,1 | 0,0 | 1,0 | 0,1 | 0,9 |
| ($e_2$) | 0,2 | 0,7 | 0,1 | 0,7 | 0,3 | 0,0 | 1,0 | 0,0 | 0,2 | 0,8 |
| ($e_3$) | 0,5 | 0,5 | 0,0 | 0,9 | 0,1 | 0,0 | 1,0 | 0,0 | 0,9 | 0,1 |
| ($e_4$) | 0,1 | 0,6 | 0,3 | 0,2 | 0,2 | 0,6 | 1,0 | 0,0 | 0,9 | 0,1 |

# Why Fuzzy Knowledge?

- People do not often use exact data at all


# Fuzzy Logics

- A tool for involving cognitive uncertainties and computing with words familiar to humans


- There are three basic terms [Yen, 1999]:
    - Fuzzy set
    - Possibility distribution
    - Linguistic variable

**[Yen, 1999]** Yen, J.: Fuzzy logic - a modern perspective. *IEEE Transactions on Knowledge and Data Engineering* 11, 1999, 153-165.

janbohacik@yahoo.com

# Fuzzy Set [Zadeh, 1965]

Let some universe ***U*** be given, a fuzzy subset ***M*** of the set ***U*** is defined with the membership function $\mu_M$ :

$$\mu_M : U \rightarrow \langle 0, 1 \rangle,$$

where the value of $\mu_M$ (***x***) for each ***x*** in ***U*** is interpreted as the degree to which the ***x*** is an element of ***M***, or equally, as the truthfulness of the statement "***x*** is an element of ***M***".

**[Zadeh, 1965]** Zadeh L.: Fuzzy set. *Journal of Information and Control* 8, 1965, 338-353.

# Fuzzy Set

**Complement** (2 axioms):

$$\text{e.g.}\quad \mu_{\bar{M}}(\boldsymbol{x}) = 1 - \mu_{\boldsymbol{M}}(\boldsymbol{x})$$

**Union** (4 axioms for so-called S-norm operator):

$$\text{e.g.}\ \mu_{\boldsymbol{M}\cup\boldsymbol{N}}(\boldsymbol{x}) = \mathbf{S}(\mu_{\boldsymbol{M}}(\boldsymbol{x}), \mu_{\boldsymbol{N}}(\boldsymbol{x})) = \boldsymbol{max}\{\mu_{\boldsymbol{M}}(\boldsymbol{x}), \mu_{\boldsymbol{N}}(\boldsymbol{x})\}$$

**Intersection** (4 axioms for so-called T-norm operator):

$$\text{e.g.}\ \mu_{\boldsymbol{M}\cap\boldsymbol{N}}(\boldsymbol{x}) = \mathbf{T}(\mu_{\boldsymbol{M}}(\boldsymbol{x}), \mu_{\boldsymbol{N}}(\boldsymbol{x})) = \boldsymbol{min}\{\mu_{\boldsymbol{M}}(\boldsymbol{x}), \mu_{\boldsymbol{N}}(\boldsymbol{x})\}$$

**Cardinality**:

$$\mathbf{M}(\boldsymbol{M}) = \sum_{x\in \boldsymbol{U}} \mu_{\boldsymbol{M}}(\boldsymbol{x})$$

**α-cut**:

$$\mu_{\boldsymbol{M}}^{\alpha}(\boldsymbol{x}) = \begin{cases} \mu_{\boldsymbol{M}}(\boldsymbol{x}); \text{ ak } \mu_{\boldsymbol{M}}(\boldsymbol{x}) \geq \alpha \\ 0 \qquad\qquad ; \text{ ak } \mu_{\boldsymbol{M}}(\boldsymbol{x}) < \alpha \end{cases}$$

# Possibility distribution

•A possibility distribution example (some modification from [Zimmernan, 1993]):

a) Probability distribution

| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
|---|---|---|---|---|---|---|---|---|---|
| $P(X_{Ján}=x_i)$ | 0.2 | 0.7 | 0.1 | 0 | 0 | 0 | 0 | 0 | $\sum = 1$ |

b) Possibility distribution

| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
|---|---|---|---|---|---|---|---|---|---|
| $\pi(X_{Ján}=x_i)$ | 1 | 1 | 1 | 1 | 0.8 | 0.6 | 0.4 | 0.2 | $\sum \geq 0$ |

**[Zimmerman, 1993]** Zimmerman, H.-J.: *Fuzzy Set Theory and its Applications 2nd Edition*, Kluwer, Boston, 1993.

# Linguistic variable

***Linguistic variable*** is a (lexical) name that is associated with a universe $\boldsymbol{U}$ and whose value may be any fuzzy subset $\boldsymbol{M}$ of the universe $\boldsymbol{U}$.

***Linguistic term*** is a (lexical) name associated with a fuzzy set $\boldsymbol{M}$ that is defined on the universe $\boldsymbol{U}$ of a linguistic variable.

$$\text{Marking: } \mu_{Variable \text{ is } term}(\boldsymbol{x}) \Leftrightarrow \mu_{term}(\boldsymbol{x}) \Leftrightarrow \mu_{\boldsymbol{M}}(\boldsymbol{x})$$

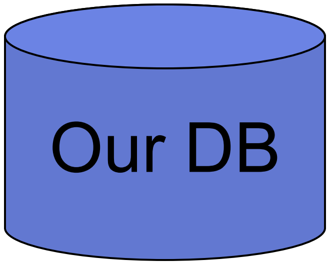
According to the formulas in [Klement and Slany, 1994]:

$$\mu_{term_1 \text{ OR } term_2}(\boldsymbol{x}) = \mathbf{S}\left(\mu_{term_1}(\boldsymbol{x}),\ \mu_{term_2}(\boldsymbol{x})\right)$$

$$\mu_{term_1 \text{ AND } term_2}(\boldsymbol{x}) = \mathbf{T}\left(\mu_{term_1}(\boldsymbol{x}),\ \mu_{term_2}(\boldsymbol{x})\right)$$

**[Klement and Slany, 1994]** Klement, E. P., Slany, W.: Fuzzy logic in artificial intelligence. CD-Technical Report 94/67, Christian Doppler Laboratory for Expert Systems E184/2, TU Wien, Wienna, Austria, 1994.

# Fuzzy Rule Base




$\mathbf{A}=\{A_1;\ \ldots\ A_k\ \ldots;\ A_n\},\ A_k=\{a_{k,1};\ \ldots;\ a_{k,l};\ \ldots;\ a_{k,n_k}\}$

$A_k$ - linguistic variable $a_{k,l}$ - linguistic term

$$\begin{array}{l}
\textbf{IF } \boldsymbol{E}_1^{Condition} \textbf{ THEN } \boldsymbol{E}_1^{Conclusion} \\
\qquad\qquad\qquad\cdot \\
\textbf{IF } \boldsymbol{E}_i^{Condition} \textbf{ THEN } \boldsymbol{E}_i^{Conclusion} \\
\qquad\qquad\qquad\cdot \\
\textbf{IF } \boldsymbol{E}_m^{Condition} \textbf{ THEN } \boldsymbol{E}_m^{Conclusion}
\end{array}$$

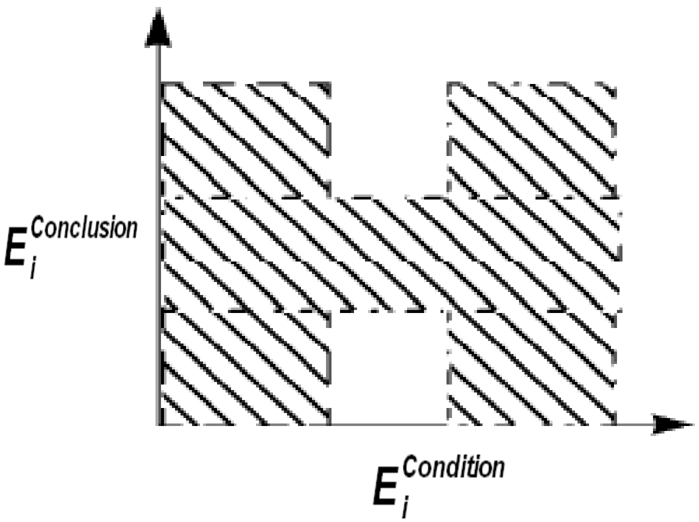
$\boldsymbol{E}_i^{Condition} \cap \boldsymbol{E}_i^{Conclusion} = \emptyset,\ \mathrm{M}(\boldsymbol{E}_i) \geq 1$

$\boldsymbol{E}_i = A_{i_1}$ is $a_{i_1}$ **AND** $A_{i_2}$ is $a_{i_2}$ **AND** ... **AND** $A_{i_{n_i}}$ is $a_{i_{n_i}}$

max one $\boldsymbol{A}_k \in \boldsymbol{E}_i^{Condition} \cup \boldsymbol{E}_i^{Conclusion}$

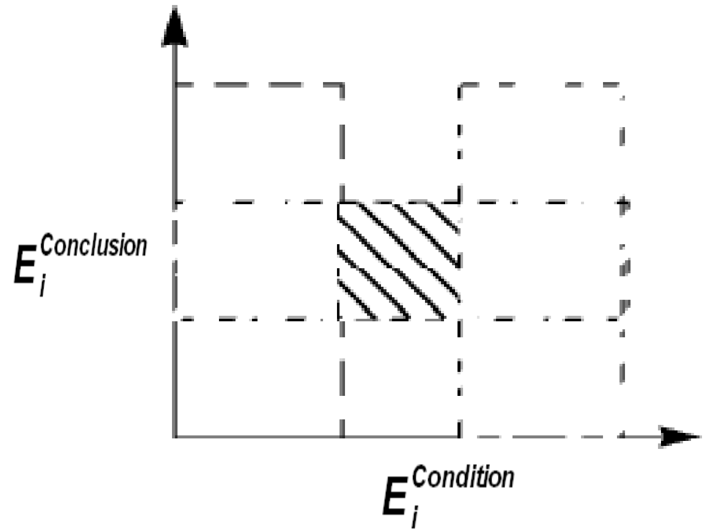
# Interpretations of a fuzzy rule [Yen, 1999]

- Implication Rule




- Designed individually
- Artificial Intelligence
- Generalized boolean implication

- Mapping Rule




- Designed as a group
- Discoverable easier
- Good for data mining

**[Yen, 1999]** Yen, J.: Fuzzy logic - a modern perspective. *IEEE Transactions on Knowledge and Data Engineering* 11, 1999, 153-165.

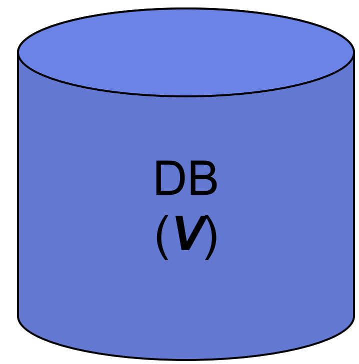
# Example of a Database with Fuzzy Data

| V | $A_1$ | | | $A_2$ | | | $A_3$ | | | $A_4$ | | $A_5$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a_{11}$ | $a_{12}$ | $a_{13}$ | $a_{21}$ | $a_{22}$ | $a_{23}$ | $a_{31}$ | $a_{32}$ | $a_{33}$ | $a_{41}$ | $a_{42}$ | $a_{51}$ | $a_{52}$ |
| $e_1$ | 0.2 | 0.7 | 0.1 | 0.3 | 0.7 | 0.0 | 0.2 | 0.8 | 0.0 | 0.0 | 1.0 | 0.4 | 0.6 |
| $e_2$ | 0.9 | 0.1 | 0.0 | 1.0 | 0.0 | 0.0 | 0.8 | 0.1 | 0.1 | 0.6 | 0.5 | 0.2 | 0.8 |
| $e_3$ | 0.8 | 0.2 | 0.0 | 0.6 | 0.4 | 0.0 | 0.0 | 1.0 | 0.0 | 1.0 | 1.0 | 0.0 | 0.7 |
| $e_4$ | 0.0 | 0.7 | 0.3 | 0.8 | 0.2 | 0.0 | 0.1 | 0.9 | 0.0 | 0.8 | 0.2 | 0.1 | 0.9 |
| $e_5$ | 0.0 | 0.1 | 0.9 | 0.7 | 0.3 | 0.0 | 0.3 | 0.4 | 0.3 | 0.5 | 0.5 | 1.0 | 0.0 |
| $e_6$ | 0.0 | 0.7 | 0.3 | 0.0 | 0.3 | 0.7 | 0.7 | 0.3 | 0.0 | 0.8 | 0.2 | 0.8 | 0.2 |
| $e_7$ | 0.9 | 0.1 | 0.0 | 0.2 | 0.8 | 0.0 | 0.1 | 0.9 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| $e_8$ | 0.0 | 0.9 | 0.1 | 0.0 | 0.9 | 0.1 | 0.1 | 0.9 | 0.0 | 0.7 | 0.0 | 1.0 | 0.0 |
| $e_9$ | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.6 | 0.0 | 0.4 | 0.8 | 0.2 | 1.0 | 0.0 |
| $e_{10}$ | 1.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| $e_{11}$ | 0.0 | 0.3 | 0.7 | 0.0 | 0.0 | 1.0 | 0.0 | 1.0 | 0.1 | 0.9 | 0.1 | 1.0 | 0.0 |
| $e_{12}$ | 0.0 | 1.0 | 0.0 | 0.0 | 0.2 | 0.8 | 0.2 | 0.8 | 0.0 | 1.0 | 0.0 | 0.7 | 0.3 |
| $e_{13}$ | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.3 | 0.0 | 0.7 | 0.6 | 0.4 | 0.2 | 0.8 |
| $e_{14}$ | 0.9 | 0.1 | 0.0 | 0.0 | 0.3 | 0.7 | 0.0 | 1.0 | 0.9 | 0.1 | 0.9 | 0.7 | 0.3 |
| $e_{15}$ | 0.7 | 0.3 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.8 | 0.2 | 0.3 | 0.7 |
| $e_{16}$ | 0.2 | 0.6 | 0.2 | 0.0 | 1.0 | 0.0 | 0.0 | 0.7 | 0.3 | 0.7 | 0.3 | 0.4 | 0.6 |

$e_i \in V \subseteq U$

***V*** – known instances

***U*** - universe




**A** = {$A_1$; $A_2$; …; $A_5$} = {*Temp(erature); Outlook*; *Bargain*; *Attendance*; *Shopping*}

*Temp* = {$a_{11}$; $a_{12}$; $a_{13}$} = {*hot*; *mild*; *cool*}

*Outlook* = {$a_{21}$; $a_{22}$; $a_{23}$} = {*sunny*; *cloudy*; *rainy*}

*Novelty* = {$a_{31}$; $a_{32}$; $a_{33}$} = {*none*; *low*; *high*}

*Attendance* = {$a_{41}$; $a_{42}$} ={*low*; *high*}

*Shopping* = {$a_{51}$; $a_{52}$} ={*good*; *bad*}

janbohacik@yahoo.com

# Classification of fuzzy rules

- Classification Fuzzy Rules

**IF** $\boldsymbol{E}_i^{Condition}$ **THEN** $C$ is $c_j$ $\qquad C \in \mathbf{A}, \; C \notin \boldsymbol{E}_i^{Condition}$

$C = A_5 =$ *Shopping*

**IF** *Outlook* is *sunny* **THEN** *Shopping* is *bad*

**IF** *Outlook* is *cloudy* **AND** *Bargain* is *none* **THEN** *Shopping* is *good*

Making humans' expected decisions on the basis of their previous decisions

- Association Fuzzy Rules

**IF** *Outlook* is *rain* **THEN** *Shopping* is *good* **AND** *Temp* is *mild*

**IF** *Shopping* is *good* **THEN** Bargain is high

- Discovering interesting associations among linguistic terms of collected instances

# FRs induction and classification

Our DB

Possibilities for terms in the class variable are being determined:

| ***U*** | ***A₁*** | | | ***A₂*** | | | ***A₃*** | | | ***A₄*** | | ***Shopping* (*C*)** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a_{1,1}$ | $a_{1,2}$ | $a_{1,3}$ | $a_{2,1}$ | $a_{2,2}$ | $a_{2,3}$ | $a_{3,1}$ | $a_{3,2}$ | $a_{3,3}$ | $a_{4,1}$ | $a_{4,2}$ | *good* ($c_1$) | *bad* ($c_2$) |
| $e_{new}$ | 0.2 | 0.7 | 0.1 | 0.3 | 0.7 | 0.0 | 0.2 | 0.8 | 0.0 | 0.0 | 1.0 | ? | ? |

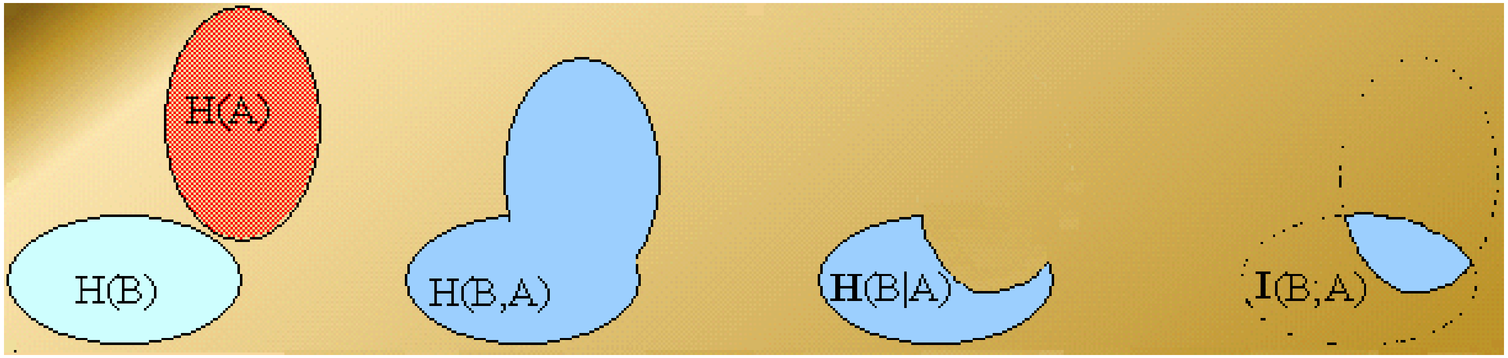
**IF** *Outlook* is *sunny* **THEN** *Shopping* is *bad*
**IF** *Outlook* is *cloudy* **AND** *Novelty* is *none* **THEN** *Shopping* is *good*
**IF** *Outlook* is *cloudy* **AND** *Temperature* is *hot* **AND** *Attendance* is *low* **THEN** *Shopping* is *bad*
**IF** *Outlook* is *cloudy* **AND** *Temperature* is *hot* **AND** *Attendance* is *high* **THEN** *Shopping* is *good*
**IF** *Outlook* is *cloudy* **AND** *Temperature* is *mild* **THEN** *Shopping* is *good*
**IF** *Temperature* is *cool* **THEN** *Shopping* is *good*
**IF** *Outlook* is *cloudy* **AND** *Novelty* is *high* **THEN** *Shopping* is *good*
**IF** *Outlook* is *rainy* **THEN** *Shopping* is *good*

janbohacik@yahoo.com

# Algorithm Based on a FDT

- Uses cumulative information [Levashenko and Zaitseva, 2002]

| | Proper | Joint | Conditional | Mutual |
|---|---|---|---|---|
| Information | $\mathbf{I}(a_{i_1,j_1})$ | $\mathbf{I}(a_{i_1,j_1}, a_{i_2,j_2})$ | $\mathbf{I}(a_{i_1,j_1} \mid a_{i_2,j_2})$ | $\mathbf{I}(a_{i_1,j_1}; a_{i_2,j_2})$ |
| Entropy | $\mathbf{H}(A_{i_1})$ | $\mathbf{H}(A_{i_1}, A_{i_2})$ | $\mathbf{H}(A_{i_1} \mid A_{i_2})$ <br> $\mathbf{H}(A_{i_1} \mid a_{i_2,j_2})$ | $\mathbf{I}(A_{i_1}; A_{i_2})$ |




**[Levashenko and Zaitseva, 2002]** Levashenko, V., Zaitseva, E.: Usage of new information estimations for induction of fuzzy decision trees. In Proceedings of the 3rd IEEE International Conference on Intelligent Data Engineering and Automated Learning, Manchester, UK, 2002, 493-499.

# Algorithm Based on a FDT

- New criteria for choosing expanded variables

Unordered FDT

$$\frac{\mathbf{I}\left(C;\ a_{i_1,j_1},\ \ldots,\ a_{i_{q-1},j_{q-1}},\ a_{i_q,j_q}\right)}{\mathbf{Cost}\left(A_{i_q}\right)} \to \boldsymbol{max}$$

Ordered FDT

$$\frac{\mathbf{I}\left(C;\ A_{i_1},\ \ldots,\ A_{i_{q-1}},\ A_{i_q}\right)}{\mathbf{Cost}\left(A_{i_q}\right)} \to \boldsymbol{max}$$

Stable FDT

$$\frac{\mathbf{I}\left(A_{i_q};\ C,\ A_{i_1},\ \ldots,\ A_{i_{q-1}}\right)}{\mathbf{Cost}\left(A_{i_q}\right)} \to \boldsymbol{max}$$

# Algorithm Based on a FDT

- Classification with weighting fuzzy rules [Levashenko et al., 2006]

$r$=1: **IF** *Outlook* is *sunny* **THEN** *Shopping* is *bad* $W_1(\boldsymbol{e}_{new})$

$r$=2: **IF** *Outlook* is *cloudy* **AND** … **THEN** *Shopping* is *good* $W_2(\boldsymbol{e}_{new})$

⋮ ⋮

$r=R$: **IF** *Outlook* is *rainy* **THEN** *Shopping* is *good* $W_R(\boldsymbol{e}_{new})$

| *U* | ***Shopping*** (***C***) | |
|---|---|---|
| | *good* ($c_1$) | *bad* ($c_2$) |
| $\boldsymbol{e}_{new}$ | 0.7499 | ? |

$$= \sum_{r=1}^{R} \text{truthfulness}_r^{c_2} * W_r(\boldsymbol{e}_{new})$$

**[Levashenko et al., 2006]** Levashenko, V., Matiasko, K., Bohacik, J., Kovalik, S.: Learning fuzzy rules from fuzzy decision trees. *Journal of Information, Control and Management Systems* 4, 2006, 143-154.

# Algorithm Based on Variable Elimination

# Algorithm Based on Variable Elimination

| | |
|---|---|
| $\{\mu_{c_j}(\boldsymbol{e})\}$ = classify(fuzzy rules; $\boldsymbol{e}$; $C$) | |
| Step 1 | Compute ${}^{j}\mu^{i}{}_{E}(\boldsymbol{e})$ for each fuzzy rule IF $\boldsymbol{E}_i$ THEN $C$ is $c_j$. |
| Step 2 | Divide the fuzzy rules into groups marked with $j$ on the basis of their conclusions $c_j \in C$. |
| Step 3 | Using s-norm operator S(a,b) = max{a, b} unite ${}^{j}\mu^{i}{}_{E}(\boldsymbol{e})$ in each group $j$. The values of these unions are values of $\mu_{c_j}(\boldsymbol{e})$. Put differently, $\mu_{c_j}(\boldsymbol{e}) = \max\{{}^{j}\mu^{i}{}_{E}(\boldsymbol{e}) \mid \forall\ i\}$. If there is not any rule in some group $j$, set $\mu_{c_j}(\boldsymbol{e}) = 0$. |

# Algorithm Based on Variable Elimination

- Why do we need another algorithm?

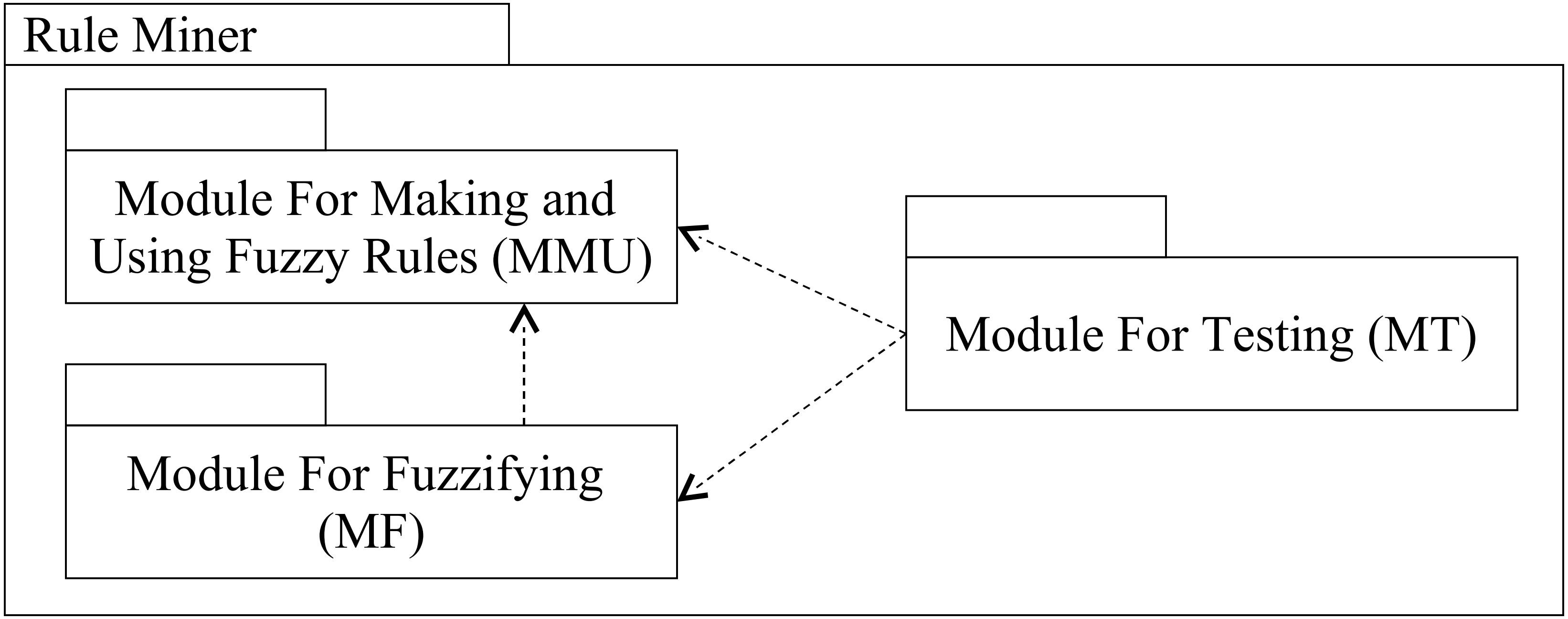
a) Fuzzy decision tree algorithms suffer from the replication and fragmentation problems [Su and Zhang, 2005]

b) Elimination unimportant variables as soon as possible may be quicker because these variables are not considered after their eliminations

c) Potential more general fuzzy rules also with OR and linguistic modifiers may lead to more accurate knowledge

**[Su and Zhang, 2005]** Su, J., Zhang, H.: Representing Conditional Independence Using Decision Trees. *In Proc. of the 12th National Conference on Artificial Intelligence*, AAAI Press, Pittsburgh, 2005, 874-879.

# Rule Miner [Bohacik et al., 2006]

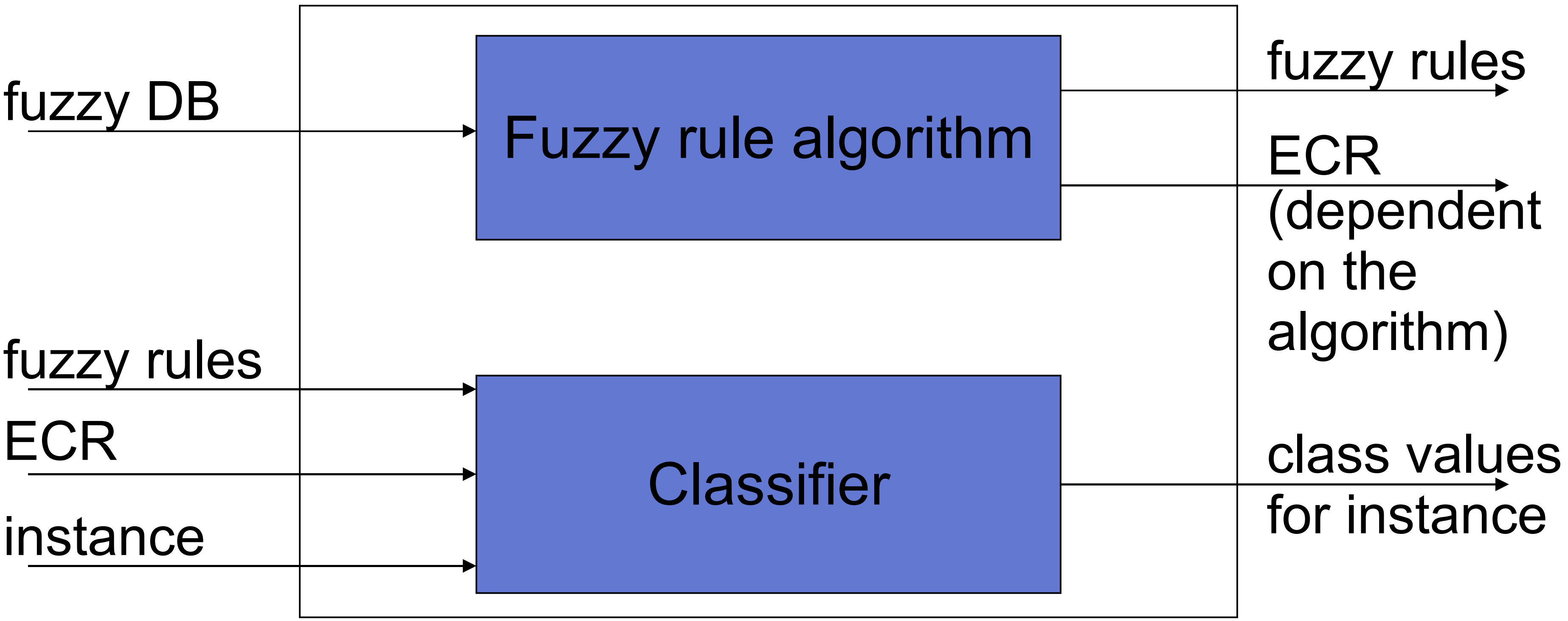



- A library written in Java (now partially in C++)
- Implements existed, modified and new algorithms
- For our research, potentially for commercial applications

**[Bohacik et al., 2006]** Bohacik, J., Matiasko, K., Levashenko, V.: Software for making and using fuzzy rules. *Journal of ELECTRICAL ENGINEERING* 57, 2006, 85-88.

janbohacik@yahoo.com

# Module For Making and Using FR (MMU)




- Implements:

**[Yuan and Shaw, 1995]** Yuan, Y., Shaw, M. J.: Induction of fuzzy decision trees. *Fuzzy Sets and Systems* 69, 1995, 125-139.

**[Levashenko et al., 2006]** Levashenko, V., Matiasko, K., Bohacik, J., Kovalik, S.: Learning fuzzy rules from fuzzy decision trees. *Journal of Information, Control and Management Systems* 4, 2006, 143-154.

**[Bohacik, 2007]** Bohacik, J.: Induction by fuzzy attribute elimination. [sent for reviewing]

janbohacik@yahoo.com

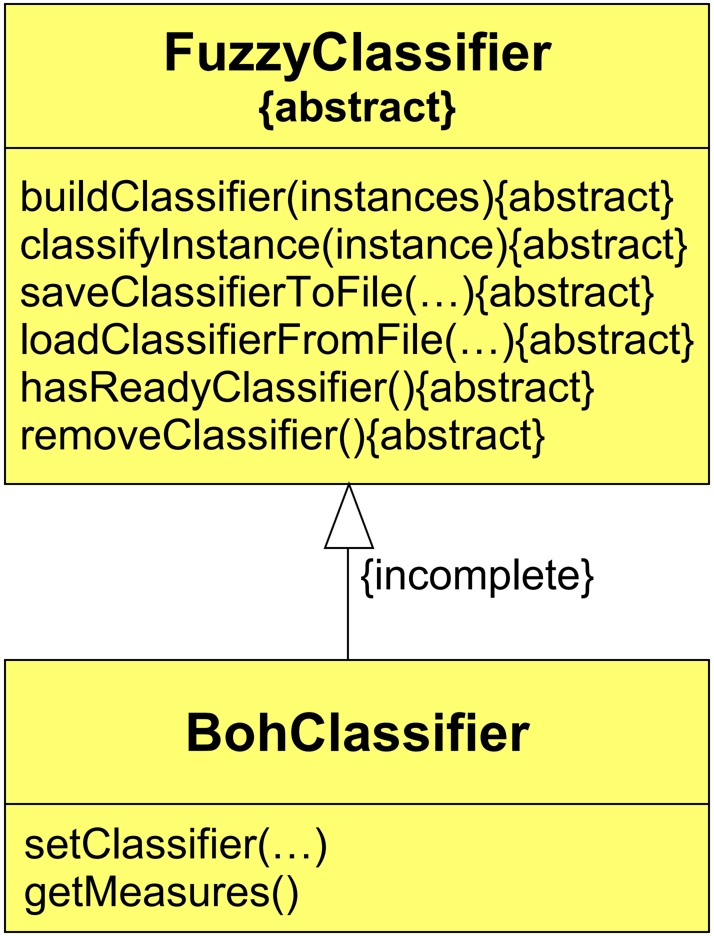
# An Example of Using MMU

```
VariablesForClassification variables = new
          VariablesForClassification(“DB.vars");
Variables.setClassVariable(“Shopping”);
InstancesForClassification instances = new
          InstancesForClassification(variables);
instances.loadFromFile(“DB.insts");
FuzzyClassifier classifier = new  BohClassifier(...);
(BohClassifier)classifier.setClassifier(0.2, 0.7);
classifier.buildClassifier();
double[] classes =
   classifier.classifyInstance(instances.getInstance(14));
System.out.println(“original classes: ” + …);
System.out.println(“made classes: ” +
           MyArrays.toString(classes));
```

**Output for instance 14:**

Original classes: 0.7 0.3

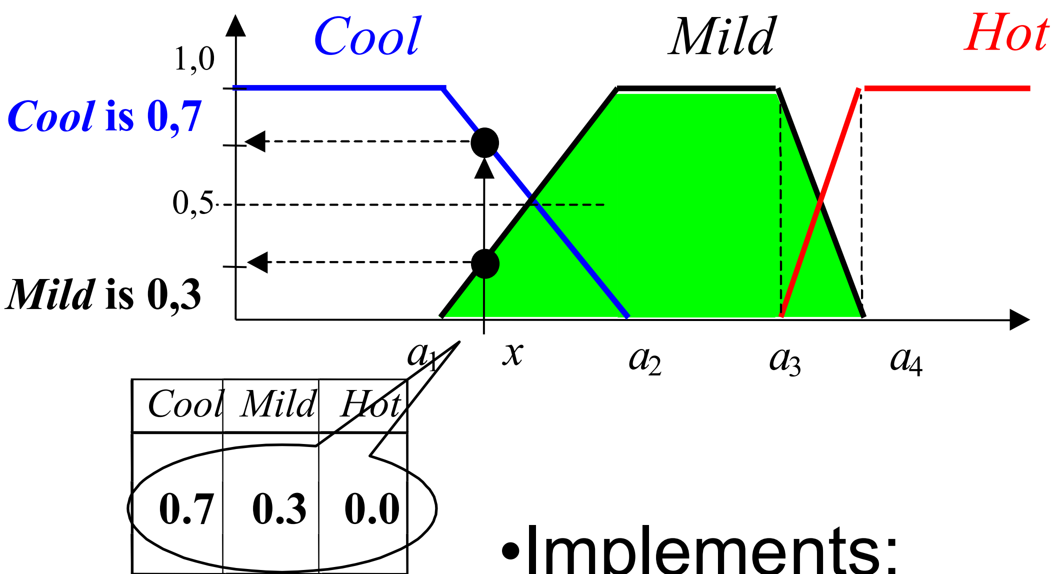
Made classes: 0.7 0.1



# Module for Fuzzifying (MF)

- Transforms numerical data into fuzzy domain
- Main problems:
  - The number of intervals
  - The membership function for each interval

$$\mu_{Cool}(x) = \begin{cases} 1, & \text{for } x < a_1 \\ \dfrac{a_2 - x}{a_2 - a_1}, & \text{for } a_1 \le x \le a_2 \\ 0, & \text{otherwise.} \end{cases}$$



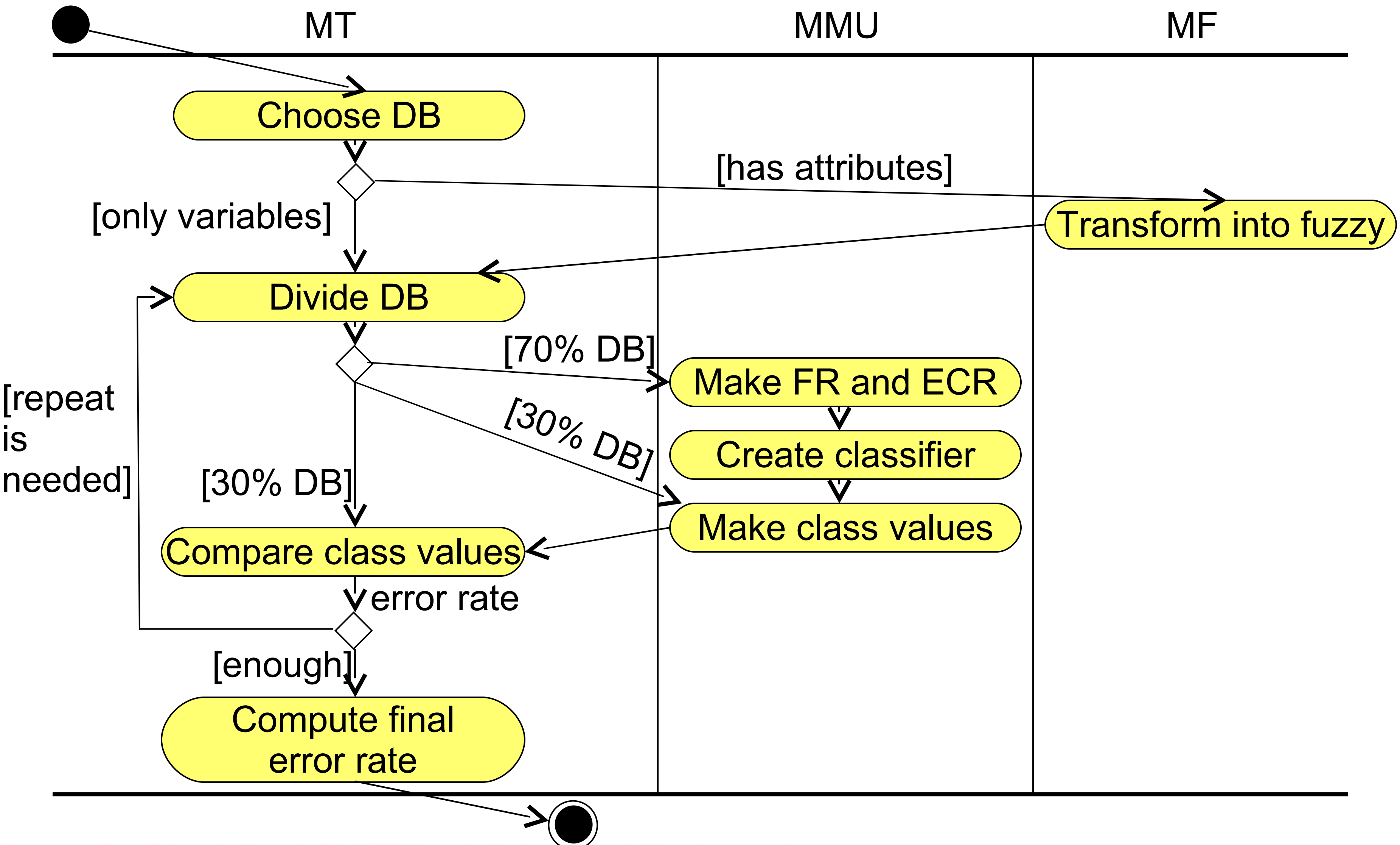

$$\mu_{Mild}(x) = \begin{cases} 0, & \text{for } x < a_1 \\ \dfrac{x - a_1}{a_2 - a_1}, & \text{for } a_1 \le x \le a_2 \\ 1, & \text{for } a_2 < x < a_3 \\ \dfrac{a_4 - x}{a_4 - a_3}, & \text{for } a_3 \le x \le a_4 \\ 0, & \text{otherwise.} \end{cases}$$

$$\mu_{Hot}(x) = \begin{cases} 0, & \text{for } x < a_3 \\ \dfrac{x - a_3}{a_4 - a_3}, & \text{for } a_3 \le x \le a_4 \\ 1, & \text{otherwise.} \end{cases}$$

- Implements:

**[Lee et al., 2001]** Lee, H.-M., Chen, C.-M., Chen, J.-M., Jou, Y.-L.: An efficient fuzzy classifier with feature selection based on fuzzy entropy. *IEEE Transaction on Systems, Man, and Cybernetics – Part B: Cybernetics* 3, 2001, 426-432.

# Module For Testing (MT)



# Algorithm Comparison

- UCI Repository of Machine Learning Databases [Hettich et al., 1998]

| Method/ DB | FAE | CI-M | CI-O | CA-MM | k-NNC | NBC |
|---|---|---|---|---|---|---|
| BUPA | 0.4214[4] | 0.4174[2] | 0.4205[3] | 0.4253[5] | 0.3910[1] | 0.4416[6] |
| Ecoli | 0.1879[2] | 0.2022[3] | 0.2654[5] | 0.3112[6] | 0.2046[4] | 0.1547[1] |
| Glass | 0.3178[1] | 0.3988[3] | 0.4647[5] | 0.4430[4] | 0.3335[2] | 0.5394[6] |
| Haberman | 0.2449[1] | 0.2624[5] | 0.2609[3] | 0.2614[4] | 0.3471[6] | 0.2453[2] |
| Iris | 0.006444[1] | 0.04067[4] | 0.02956[2] | 0.04000[3] | 0.05044[6] | 0.04556[5] |
| Pima | 0.2520[4] | 0.2436[1] | 0.2563[5] | 0.2509[3] | 0.3091[6] | 0.2483[2] |
| Wine | 0.04660[3] | 0.04566[2] | 0.06509[5] | 0.08170[6] | 0.05094[4] | 0.02697[1] |
| **Average** | 0.2110 [2.2875] | 0.2301 [2.851] | 0.2518 [4.0000] | 0.2591 [4.4286] | 0.2409 [4.1429] | 0.2431 [3.2875] |

**[Hettich et al., 1998]** Hettich, S., Blake C. L., Merz, C. J.: UCI Repository for Machine Learnig Data-Bases, University of California, Irvine, 1998, on http://www.ics.uci.edu/~mlearn/MLRepository.html.

janbohacik@yahoo.com

# Discussion

janbohacik@yahoo.com